

ON PARAMETERIZED COMPLEXITY OF THE HITTING SET PROBLEM FOR AXIS-PARALLEL SQUARES INTERSECTING A STRAIGHT LINE

Daniel M. Khachay

N.N. Krasovskii Institute of Mathematics and Mechanics,
Ural Branch of the Russian Academy of Sciences and
Ural Federal University, Ekaterinburg, Russia,
dmx@imm.uran.ru

Michael Yu. Khachay

N.N. Krasovskii Institute of Mathematics and Mechanics,
Ural Branch of the Russian Academy of Sciences and
Ural Federal University, Ekaterinburg, Russia,
mkhachay@imm.uran.ru

Abstract: The Hitting Set Problem (HSP) is the well known extremal problem adopting research interest in the fields of combinatorial optimization, computational geometry, and statistical learning theory for decades. In the general setting, the problem is NP-hard and hardly approximable. Also, the HSP remains intractable even in very specific geometric settings, e.g. for axis-parallel rectangles intersecting a given straight line. Recently, for the special case of the problem, where all the rectangles are unit squares, a polynomial but very time consuming optimal algorithm was proposed. We improve this algorithm to decrease its complexity bound more than 100 degrees of magnitude. Also, we extend it to the more general case of the problem and show that the geometric HSP for axis-parallel (not necessarily unit) squares intersected by a line is polynomially solvable for any fixed range of squares to hit.

Key words: Hitting set problem, Dynamic programming, Computational geometry, Parameterized complexity.

Introduction

We consider the parameterized complexity of a geometric statement of the well-known Hitting Set Problem (HSP), engaging researchers in combinatorial optimization, computational geometry and statistical learning from early 1980-th.

To the best of our knowledge, HSP gains theoretical interest because it was the first intractable combinatorial optimization problem, whose approximation algorithms were dramatically improved [11] on the basis of Vapnik and Chervonenkis's [15] results in statistical learning theory. The development of randomized algorithms for HSP and related combinatorial problems defined on range spaces of finite VC-dimension, initiated by seminal papers [1] and [6] established a new field in modern computational geometry.

On the other hand, the concepts of hitting set and classifier ensemble, making decisions by some voting logic, seem to be related very closely. Consequently, approximation techniques developed for HSP and its dual Set Cover problem are closely related to the well-known boosting learning technique [14], especially in the context of the minimal committee problem looking for minimum VC-dimension correct majority classifier ensemble (see, e.g., [8–10]).

In addition, new efficient optimal and approximation algorithms for Hitting Set and Set Cover problems have a practical importance, e.g. in design of reliable wireless networks [13].

The Hitting Set Problem for Axis-Parallel Rectangles (HSP-APR) is a well-studied geometric setting of the HSP. This setting is also NP-hard [5] and remains intractable even for unit squares.

In papers [2, 7], first polynomial time approximation schemes (PTAS) are proposed for axis-parallel squares. Paper [3] introduces 6-approximation polynomial time algorithm for the case of rectangles intersecting some axis-monotone curve. In [4], this particular case of HSP–APR is proved to be NP-hard even for a straight line and the first 4-approximation algorithm is constructed.

In this paper, we improve one of the recent results describing a polynomial time solvable subclass of this problem. Recently, Mudgal and Pandit [12] introduced an optimal polynomial time algorithm for the Hitting Set Problem for Axis Parallel Unit Squares Intersecting a given Straight Line (HSP–APUS–ISL). The theoretical importance of this result can hardly be overestimated, since almost all known geometric settings of the HSP, including extremely specific ones, are intractable. Unfortunately, this algorithm is impractical due to its incredibly high time consumption of $O(n^{145})$. In Section 2, we propose the improved version of the algorithm, whose complexity bound $O(n^{37})$ is still high but by more than 100 degrees of magnitude better. Further, in Section 3, we extend this algorithm on a case of squares of different sizes (HSP–APS–ISL) and show that this problem can be solved to optimal in polynomial time for any fixed range of square sizes.

1. Problem statement

We consider the following geometric setting of the well-known Hitting Set Problem, which is called the Hitting Set Problem for Axis-Parallel Squares Intersecting a Straight Line (HSP–APS–ISL) (see Fig. 1). In the Euclidean plane, a finite collection $S = \{Q_1, \dots, Q_n\}$ of axis-parallel (closed) squares intersecting some straight line d is given. For the collection S , it is required to find a hitting set P^* of the minimum size, i.e.

$$P^* = \arg \min\{|P| : P \subset \mathbb{R}^2, P \cap Q_j \neq \emptyset, j = 1, \dots, n\}.$$

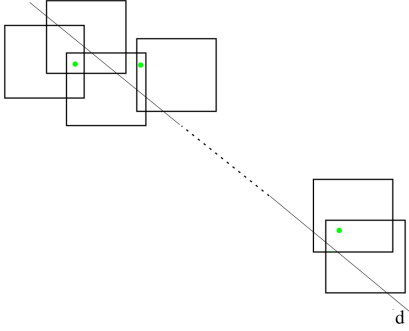


Figure 1. Problem statement

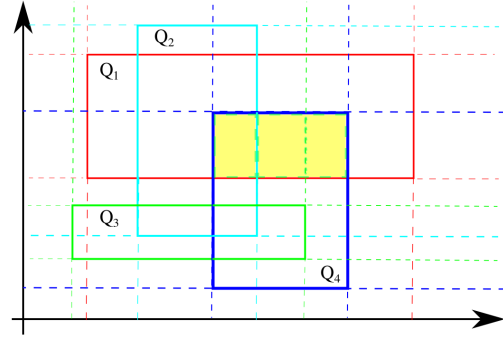


Figure 2. K does not exceed the number of rectangular cells induced by the lines defining borders of Q_1, \dots, Q_n

Without loss of generality we assume that the line d is defined by the equation $kx + y = 0$ for some $k \geq 1$.

The collection S partitions the plane onto mutually disjunctive regions $\theta_1, \dots, \theta_K$ such that, any points p_1 and p_2 belong to the same region θ_k if and only if

$$(\forall Q_j \in S) ((p_1 \in Q_j) \iff (p_2 \in Q_j)).$$

Since each minimal hitting set contains at most one point p_k taken from any region θ_k , the initial continuous problem is polynomially equivalent to the corresponding combinatorial one, which is of

finding a minimal hitting set among subsets of the finite set

$$\mathcal{P} = \{p_1, \dots, p_K\}, \quad p_k \in \theta_k \setminus \bigcup_{l \neq k} \theta_l.$$

Indeed, for any collection of n axis-parallel squares (and even rectangles), the corresponding set \mathcal{P} contains at most $O(n^2)$ elements (see Fig. 2) and can be constructed in polynomial time.

2. Improved algorithm for unit squares

In this section we describe parameterized optimal algorithm for HSP–APS–ISL and discuss its application to solving the special case of this problem, HSP–APUS–ISL, where collection S consists of equal squares (without loss of generality, which are assumed to be unit). We start with the similar (but not the same) notation to introduced in [12].

First, we partition the plane by straight lines l_0, \dots, l_{r+2} orthogonal to d with distance of $\sqrt{2}/2$ between each neighboring lines such that, for each square $Q_j \in S$, its center C_j is located between l_1 and l_{r+1} (hereinafter all tightes are broken arbitrarily). For any $i = 0, \dots, r+1$, we denote by R_i the stripe located between l_i and l_{i+1} . Next, we introduce the notation $S_i = \{Q_j : Q_j \cap R_i \neq \emptyset\}$, $S_i^{in} = \{Q_j \in S_i : C_j \in R_i\}$, and $S_i^{out} = S_i \setminus S_i^{in}$. By construction, $S_i^{out} \subset S_{i-1}^{in} \cup S_{i+1}^{in}$.

As in [12], we assume that any stripe R_i is intersected at least by a single square Q_j . Further, we find an optimal hitting set recursively, by the dynamic programming procedure presented in Algorithm 1.

Algorithm 1 Parameterized exact DP based algorithm

Input: a collection $S = \{Q_1, \dots, Q_n\}$ of axis-parallel squares intersecting a given straight line d

Outer parameter: an upper bound q of the size of subsets to search for

Output: the minimum size hitting set P for S .

- 1: Construct a set \mathcal{P} induced by the collection S ; let $\mathcal{P}_i = \mathcal{P} \cap R_i$;
- 2: **for all** $U \subset \mathcal{P}_{r-1}$ and $V \subset \mathcal{P}_r$, s.t. $|U|, |V| \leq q$ **do**
- 3: define $\mathcal{W}_r = \{W \subset \mathcal{P}_{r+1} : |W| \leq q, U \cup V \cup W \cap Q_j \neq \emptyset (Q_j \in S_r)\}$ and

$$T(r, U, V) = \begin{cases} \min\{|U \cup V \cup W| : W \in \mathcal{W}_r\}, & \text{if } \mathcal{W}_r \neq \emptyset, \\ +\infty, & \text{otherwise} \end{cases}$$

- 4: **end for**
- 5: **for all** $1 \leq i \leq r-1$ **do**
- 6: **for all** $U \subset \mathcal{P}_{i-1}$ and $V \subset \mathcal{P}_i$, s.t. $|U|, |V| \leq q$ **do**
- 7: define $\mathcal{W}_i = \{W \subset \mathcal{P}_{i+1} : |W| \leq q, U \cup V \cup W \cap Q_j \neq \emptyset (Q_j \in \bigcup_{l \geq i} S_l)\}$ and

$$T(i, U, V) = \begin{cases} |U| + \min\{T(i+1, V, W) : W \in \mathcal{W}_i\}, & \text{if } \mathcal{W}_i \neq \emptyset, \\ +\infty, & \text{otherwise} \end{cases}$$

- 8: **end for**
- 9: **end for**
- 10: Output

$$\arg \min\{T(1, U, V) : U \subset \mathcal{P}_0, V \subset \mathcal{P}_1, |U|, |V| \leq q\}.$$

Indeed, for any $i = 1, \dots, r$, denote $\mathcal{P}_i = \mathcal{P} \cap R_i$. Let, for $U \subset \mathcal{P}_{i-1}$ and $V \subset \mathcal{P}_i$, $T(i, U, V)$ be the size of a smallest hitting set P for $\bigcup_{l \geq i} S_l$ such that $P \cap \mathcal{P}_{i-1} = U$ and $P \cap \mathcal{P}_i = V$. Similarly to [12], we express $T(i, U, V)$ in terms of $T(i+1, U', V')$ but for a substantially smaller subsets U' and V' .

Algorithm 1 has an outer parameter q , which meaning is twofold. On the first hand, q depends on size-length of the squares to hit and provides a uniform upper bound for the smallest size of a hitting set for an arbitrary S_i . On the other hand, q bounds the number of subset enumerated at each iteration of Algorithm 1. Therefore, its complexity bound can be defined in terms of q again.

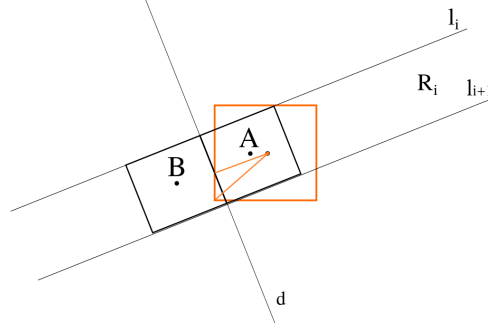


Figure 3. Any unit square $Q_j \in S_i^{in}$ is hit by one of the centers A and B of $\sqrt{2}/2$ -squares

The following Theorem summarizes the properties of Algorithm 1.

Theorem 1. *For $q = 6$, Algorithm 1 finds an optimal hitting set for the collection S in time of $O(n^{37})$.*

P r o o f. We start with the following simple fact. By construction, for any $i \in \{1, \dots, r\}$ and any $j \in S_i^{in}$, $Q_j \cap \{A, B\} \neq \emptyset$ (see Fig. 3). As a consequence, for any optimal hitting set P and any $i \in \{1, \dots, r\}$, $|P_i| \leq 6$, where $P_i = P \cap R_i$. Indeed, assume by contradiction that, for some i , $|P_i| > 6$. Since $S_i \subset S_{i-1}^{in} \cup S_i^{in} \cup S_{i+1}^{in}$ and $P_i \cap Q_j = \emptyset$ for any $Q_j \notin S_i$, we can substitute P_i by an appropriate 6-point subset P'_i such that $P \cup P'_i \setminus P_i$ remains a hitting set for S and $|P'| < |P|$. The contradiction obtained with optimality of P finalizes our argument. Hence, Algorithm 1 realizing classic dynamic programming technique finds an optimal hitting set for the given collection S .

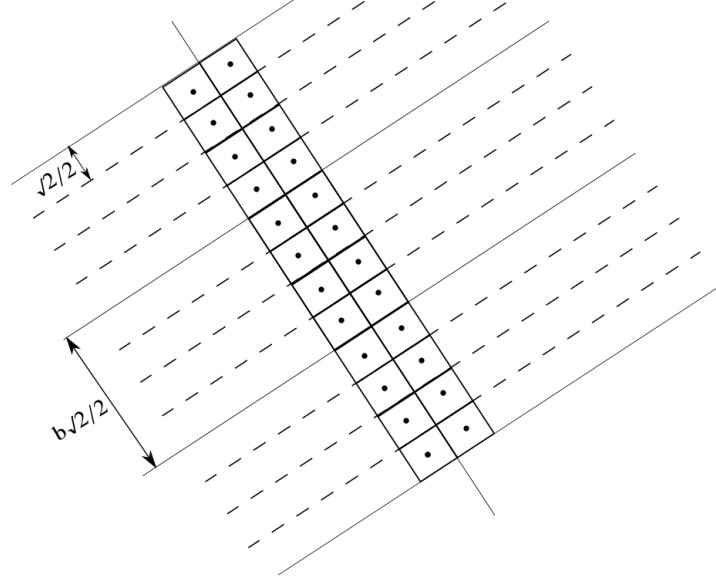
Let us obtain an upper bound for its running time. Obviously, the loop 5-9 having $r - 1 = O(n)$ iterations is the most time consuming part of Algorithm 1. In each iteration, $O(|P_{i-1}|^6) \times O(|P_i|^6) = O(n^{24})$ subproblems each having time complexity of $O(n^{12})$ should be solved. Therefore, the overall running time is $O(n^{37})$. \square

3. General case of HSP–APS–ISL

By scaling, we can easily show that the result of Section 2 remains valid in the case of equal squares of any side-length. In this section, we extend this result to the more general case. Let a and b be the minimum and the maximum values of side-lengths of the given squares. By the same reason, assume that $a = 1$.

3.1. Case of $k = 1$

We proceed with the following observation. For $k = 1$, as in Section 2, any square Q of size at least 1, whose center belongs to some stripe R' of width $\sqrt{2}/2$ orthogonal to the line d , is hit by the points A and B (like in Fig. 3). Therefore, in this case, we can adapt Algorithm 1 to take into account the squares, whose side-lengths are greater then 1.


 Figure 4. Partition of the plane for $b = 4$

Indeed, as above, consider stripes R_i of width $b\sqrt{2}/2$ consisting all the squares. Then, partition each of them onto $\lceil b \rceil$ substripes of width $\sqrt{2}/2$ (see Fig. 4) and use all other notation introduced in Section 2 as is. The following assertion is valid.

Theorem 2. *Let the given collection S consists of squares with side-lengths from $[1, b]$. Algorithm 1 with $q = 6\lceil b \rceil$ finds an optimal hitting set for this collection in time of $O(n^{6q+1}) = O(n^{36\lceil b \rceil+1})$.*

The argument proving Theorem 2 is similar to the proof of the Theorem 1. For the sake of brevity, we skip the proof.

3.2. What if $k > 1$

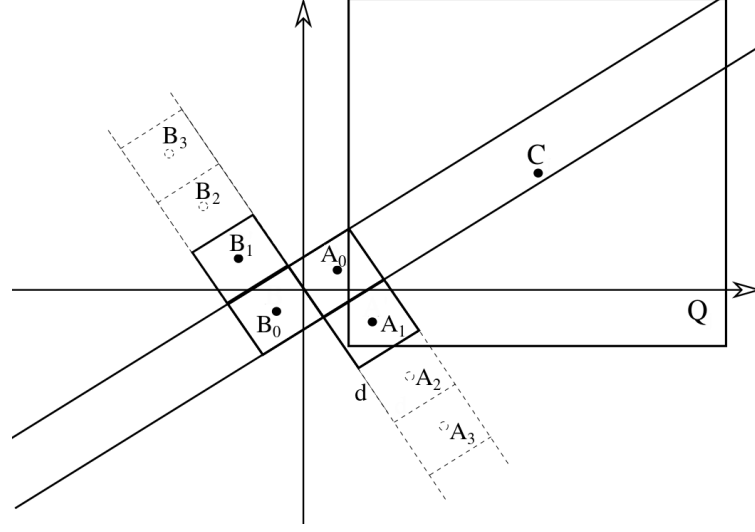
In this section, we show that to find an optimal solution for HSP–APS–ISL we can use Algorithm 1 again with an adjusted value of the parameter q . As above, this value is defined by the number of points needed to hit any square intersecting the line d , whose center belong to some stripe of the width $\sqrt{2}/2$. Although, for $k > 1$, points A and B (as in Fig. 3) do not hit all such squares, we can still provide a finite point collection that does.

Without loss of generality, assume that the strip R (of width $\sqrt{2}/2$) orthogonal to the line d is located symmetrically with respect to the origin. An arbitrary square Q intersecting the line d , whose center C belongs to the stripe R is called *R -centered*.

Consider finite point sequences $\{A_t\}$ and $\{B_t\}$ defined by the following equation

$$A_t = -B_t = \left[\frac{k + 2t}{2\sqrt{2(1+k^2)}}, \frac{1 - 2tk}{2\sqrt{2(1+k^2)}} \right] \quad (t \in \{-1, \dots, p\}). \quad (3.1)$$

Theorem 3. *For any $k > 1$, any R -centered square Q of size belonging to the range $[1, p\sqrt{2}]$ is hit by the points $A_0, \dots, A_p, B_0, B_1, \dots, B_p$.*

Figure 5. Hitting of large squares by the centers of neighboring $\sqrt{2}/2$ -squares

P r o o f. 1. Consider an arbitrary R -centered square Q . Theorem 3 is evidently valid if the center C of this square belongs to one of $\sqrt{2}/2$ -squares centered at A_0 or B_0 . Consider the other option. Without loss of generality, assume that C belongs to right-upper part of the stripe R (as in Fig. 4). The square Q coincides with an intersection of four closed halfplanes bordering it from the left, top, right, and bottom sides. We denote them by H_L, H_T, H_R , and H_B , respectively. To proceed with the argument, it is sufficient to prove that there exists a point $A_t \in Q = H_L \cap H_T \cap H_R \cap H_B$.

The inclusion $A_t \in H_T$ is valid for any $t = 0, 1, \dots, p$, since $y_{A_t} \leq y_C$ by the location assumption for the square Q . Furthermore, this assumption implies that A_{-1} can not be located to the right of the border of H_L . Suppose, $A_{t-1} \notin H_L$ and $A_i \in H_L$ for any $i \geq t$. Now, we show that A_t is the desired point hitting the square Q . Indeed, consider the intersection point D of the line d with the vertical line visiting the point A_{i-1} . Since

$$x_D = \frac{k + 2(t-1)}{2\sqrt{2}(1+k^2)}$$

and

$$kx_D + y_D = 0,$$

we obtain

$$y_{A_t} - y_D = \frac{1 - 2tk + k(k + 2(t-1))}{2\sqrt{2}(1+k^2)} = \frac{(k-1)^2}{2\sqrt{2}(1+k^2)} \geq 0.$$

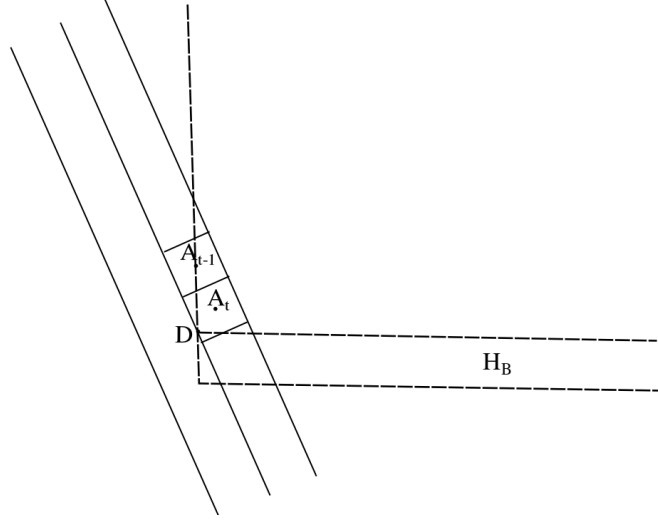
Therefore, $A_t \in H_B$ (see Fig. 6).

Inclusion $A_t \in H_R$ follows easily from equation (3.1). Indeed, for any $k > 1$

$$x_{A_t} - x_{A_{t-1}} = \frac{1}{2\sqrt{2}(1+k^2)} < 1/2 \leq x_C - x_{A_{t-1}},$$

since a size of the square Q is at least 1. Thus, $A_t \in H_L \cap H_T \cap H_R \cap H_B = Q$.

2. To obtain the upper side-length bound of the fittable squares, it is sufficient to calculate the minimum side-length of the R -centered square touching the point A_p by its left side (Fig. 7). It


 Figure 6. A_t belongs to H_B .

is easy to show that this length coincides with $s = 2(x_F - x_{A_p})$, where X_F can be found from the following system

$$\begin{cases} x_E = x_{A_p} = \frac{k + 2p}{2\sqrt{2(1+k^2)}}, \\ kx_E + y_E = 0, \\ -x_E + y_E = z, \\ -x_F + ky_F = -\frac{\sqrt{1+k^2}}{2\sqrt{2}}, \\ -x_F + y_F = z, \end{cases}$$

i.e.

$$x_F = \frac{k^3 + 2pk^2 + 2pk - 1}{2\sqrt{2}(k-1)\sqrt{1+k^2}}$$

and

$$s = \frac{k^3 + 2pk^2 + 2pk - 1}{(k-1)\sqrt{2(1+k^2)}} - \frac{k + 2p}{\sqrt{2(1+k^2)}} = \frac{\sqrt{2(1+k^2)}}{2} + \frac{p\sqrt{2(1+k^2)}}{k-1}.$$

To complete our proof, we should minimize $s = s(k)$ for $k > 1$.

The derivative

$$s'(k) = \frac{\sqrt{2} k(k-1)^2 - 2p(k+1)}{2(k-1)^2\sqrt{1+k^2}}$$

is vanishing if and only if

$$k^3 - 2k^2 + k = 2p(k+1). \quad (3.2)$$

For $p = 0$, the function $s(k)$ has no minimizers in $(1, \infty)$. The right limit

$$\lim_{k \rightarrow +0} s(k) = \inf\{s(k) : k > 1\} = 1,$$

although $s(1) = +\infty$, as it follows from Subsection 3.1.

Given that $p \geq 1$, it is sufficient to consider a few cases. If $p = 1$ we have a single root (in the feasible domain $\{k: k > 1\}$) and it is easy to see that this root is a minimizer of $s(k)$, since $s'(k)$ changes its sign at this point. Further, it can be verified that, for any $p > 1$, we also have the unique extremal point.

Denote by $\bar{k} = \bar{k}(p)$ this extremum for the given p . Using equation (3.2), we obtain

$$s(\bar{k}) = \frac{\sqrt{2}(1 + \bar{k}^2)^{3/2}}{2(1 + \bar{k})}.$$

Therefore, since $\bar{k} > 1$,

$$\frac{s(\bar{k}(p))}{p} = \frac{\sqrt{2}(1 + \bar{k}^2)^{3/2}}{\bar{k}(1 - \bar{k})^2} \geq \frac{\sqrt{2}(3/2 + \bar{k}^2)}{(\bar{k} - 1)^2} > \sqrt{2}.$$

Theorem is proved. □

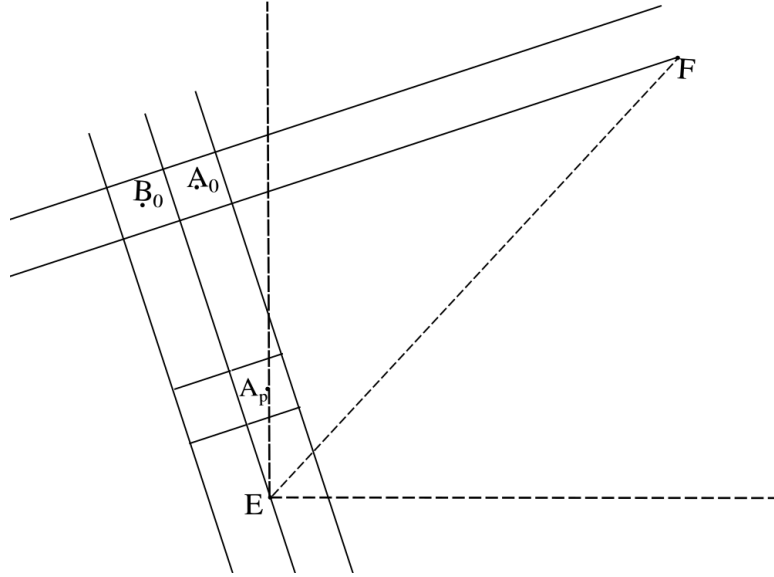


Figure 7. Estimation of $s(\bar{k})$.

Remark 1. It is easily to verify that $\bar{k} = \bar{k}(p)$ is a monotonically increasing function and tends to $+\infty$ as $p \rightarrow +\infty$. Therefore,

$$\lim_{p \rightarrow +\infty} \frac{s(\bar{k}(p))}{p} = \lim_{\bar{k} \rightarrow +\infty} \frac{\sqrt{2}(1 + \bar{k}^2)^{3/2}}{2(1 + \bar{k})} = \sqrt{2}.$$

Applying the approach proposed in Subsection 3.1, we obtain our final result. Indeed, let we should find the minimum hitting set for n squares intersecting the line d ; sizes of the squares belong to $[a', b']$. First, by scaling, transform their sizes to the range $[1, b]$, where $b = b'/a'$.

Further, partition the plane onto d -orthogonal stripes of width $b\sqrt{2}/2$; we call these stripes *wide*. Finally, we partition each wide stripe onto $\lceil b \rceil \sqrt{2}/2$ -width *narrow* substripes.

By construction, any square intersecting a wide stripe is centered at this or two neighboring wide stripes. Therefore, by Theorem 3, it can be hit by $q = 6\lceil b \rceil + 2\lceil b/\sqrt{2} \rceil$, and the optimal hitting set can be found by Algorithm 1 using this value of q . Hence, we proved the following theorem.

Theorem 4. *For any constant c and any square collection with size-range $[a, c\bar{a}]$, the problem HSP-ASP-ISL can be solved to optimality in time $O(n^{6q+1})$, where $q = 6\lceil c \rceil + 2\lceil c/\sqrt{2} \rceil$.*

Remark 2. *Results of Theorem 2 and 3 shows that HSP-APS-ISL is polynomial solvable for any fixed range of squares, since the running time bound of Algorithm 1 in this case is*

$$O(n^{6(6\lceil c \rceil + 2\lceil c/\sqrt{2} \rceil) + 1}).$$

Unfortunately, the question of constructing for this problem an FPT algorithm having parameterized complexity bound like $f(c) \cdot n^{O(1)}$ still remains open.

4. Conclusion

In the paper, the improved version of the optimal polynomial time hitting set construction algorithm for axis-parallel squares intersecting the given straight line introduced in [12] is proposed. Our modification has better upper time complexity bound by 100 orders of magnitude.

Also, we propose an extension of this algorithm to the case of non-unit squares and show that the problem can be solved to optimality in polynomial time for any fixed range of squares.

As for the future work, it would be interesting to establish the complexity status of the considered problem in the case, where this parameter is unbounded. Also, it is interesting to answer the question, does the Hitting Set Problem for Axis-Parallel Squares belong to the class of Fixed Parameter Tractable (FPT) problems.

Acknowledgements

This research was supported by Russian Foundation for Basic Research, grant no. 16-07-00266 and Complex Program of Ural Branch of RAS, grant no. 15-7-1-23.

REFERENCES

1. **Brönnimann H. and Goodrich M. T.** Almost optimal set covers in finite vc-dimension // Discrete & Computational Geometry, 1995. Vol. 14, no. 4. P. 463–479. DOI: 10.1007/BF02570718
2. **Chan T. M.** Polynomial-time approximation schemes for packing and piercing fat objects // J. of Algorithms, 2003. Vol. 46, no. 2. P.178–189. DOI: 10.1016/S0196-6774(02)00294-8
3. **Chepoi V. and Felsner S.** Approximating hitting sets of axis-parallel rectangles intersecting a monotone curve. Computational Geometry, 2013. Vol. 46, no. 9. P. 1036–1041. DOI: 10.1016/j.comgeo.2013.05.008
4. **Correa J., Feuilloley L., Perez-Lantero P. and Soto J. A.** Independent and hitting sets of rectangles intersecting a diagonal line // Algorithms and complexity. Discrete & Computational Geometry, 2015. Vol. 53, no. 2. P. 344–365. DOI: 10.1007/s00454-014-9661-y
5. **Fowler R. J., Paterson M. S. and Tanimoto S. L.** Optimal packing and covering in the plane are np-complete. Information Processing Letters, 1981. Vol. 12, no. 3. P. 133–137. DOI: 10.1016/0020-0190(81)90111-3
6. **Haussler D. and Welzl E.** Epsilon-nets and simplex range queries // Discrete & Computational Geometry, 1987. Vol. 2, no. 2. P. 127–151. DOI: 10.1007/BF02187876
7. **Hochbaum D. and Maass W.** Approximation schemes for covering and packing problems in image processing and vlsi // J. ACM, 1985. Vol. 32, no. 1. P. 130–136. DOI: 10.1145/2455.214106
8. **Khachay M.** Committee polyhedral separability: complexity and polynomial approximation // Machine Learning, 2015. Vol. 101, no. 1. P. 231–251. DOI: 10.1007/s10994-015-5505-0
9. **Khachay M. and Poberii M.** Complexity and approximability of committee polyhedral separability of sets in general position. Informatica, 2009. Vol. 20, no. 2. P. 217–234.

10. **Khachay M., Pobery M. and Khachay D.** Integer partition problem: Theoretical approach to improving accuracy of classifier ensembles // International J. of Artificial Intelligence, 2015. Vol. 13, no. 1. P. 135–146.
11. **Matoušek J.** Lectures on Discrete Geometry. Springer: New York, 2002. DOI: 10.1007/978-1-4613-0039-7
12. **Mudgal A. and Pandit S.** Covering, hitting, piercing and packing rectangles intersecting an inclined line // Proceedings of the Combinatorial Optimization and Applications: 9th International Conference, (COCOA 2015, Houston, TX, USA, December 18–20, 2015), Zaixin Lu, Donghyun Kim, Weili Wu, Wei Li, and Ding-Zhu Du (Ed.). LNCS, Springer International Publishing: Cham, 2015. Vol. 9486. P. 126–137. DOI: 10.1007/978-3-319-26626-8_10
13. **Ramakrishnan S. and Emary I. M. M. El.** Wireless sensor networks: from theory to applications. CRCPress, Taylor & Francis, 2014.
14. **Schapire R. and Freund Y.** Boosting: Foundations and algorithms. MIT Press, 2012.
15. **Vapnik V. and Chervonenkis A.** On the uniform convergence of relative frequencies of events to their probabilities // Theory Probab. Appl., 1971. Vol. 16. P. 264–280. DOI: 10.1137/1116025